

AUTOMATED EVALUATION OF BREAST CANCER DETECTION USING SVM CLASSIFIER

BASAVARAJ HIREMATH¹ & SC PRASANNAKUMAR²

¹Assistant Professor, MSRIT, Bangalore, Karnataka, India

¹Research Scholar, Jain University, Bangalore, Karnataka, India

²Professor & HOD, Department of Instrumentation Technology, R. V. College of Engineering,
Bangalore, Karnataka, India

ABSTRACT

Breast cancer in females is the most common cancer diseases and leading cause of death. In the recent years, Computer Aided Diagnosis (CAD) is very useful for detection of breast cancer. Mammography can be used as an efficient tool for breast cancer diagnosis. A computer based diagnosis and classification system can reduce unnecessary biopsy. This paper presents the tumor detection algorithm from mammogram, this study shows the outcome of applying image processing morphological operation on mammogram breast cancer image, Since micro calcification clusters are primary indicators of malignant types of breast cancer, its detection is important to prevent and treat the disease. This paper proposes a method for detection of micro calcification clusters in mammograms using sequential Difference of Gaussian filters (DoG), and Gaussian filters. These regions are classified by SVM classifier using the most dominant features which are extracted from CSLBP features and DoG features. The proposed method was tested on 75 mammographic images, from the mini-MIAS database. The methodology achieved a accuracy of 89.33%.

KEYWORDS: Region of Interest, Difference of Gaussian, Gaussian, Center Symmetric Local Binary Pattern, SVM Classifier

INTRODUCTION

Cancer is the leading cause of death worldwide and accounts for 7.6 million deaths. Deaths from cancer are expected to rise to over 11 million in 2030. Breast cancer in females is the most frequently occurring cancer diseases. The imaging techniques frequently used for the detection of breast cancer is Mammography and Magnetic Resonance Imaging. Mammography is useful in discovering tumors too small to be felt. Computerized methods are being developed to help radiologists as second opinion for the detection of abnormality in mammograms.

Mammography is the process of using low-energy X-rays (usually around 30 kVp) to examine the human breast and is used as a diagnostic and a screening tool. The goal of mammography is the early detection of breast cancer, typically through detection of characteristic masses and/or microcalcification. A mammographic image is characterized by a high spatial resolution which is adequate enough to detect subtle fine-scale signs such as microcalcification. Breast abnormality is associated with calcification and masses. Mammogram breast cancer images have the ability to assist physicians in detecting disease caused by cells' abnormal growth.

Numerous techniques have been proposed for early detection of breast cancer using mammography. The study of Neural Networks [NN] [1] detects and locates early breast cancer using a simple feed-forward back-propagation neural network. The Computer-aided diagnosis systems have been developed based on parameters extracted from micro calcification [3]. This method presents an automatic micro calcification segmentation method, based on Otsu's method and morphological filters. In the same way numbers of techniques have been proposed for early detection of breast cancer using Magnetic Resonance Image [12]. An integrated classifier that is used in mammogram Magnetic Resonance image for classification of breast cancers and abnormalities using a Multi-stage classifier is presented in this method.

In our method SVM classifier is used for early detection of breast cancer using mammography because, the result of Support Vector Machine with sigmoid kernel shows higher classification performance than other classifier. As we go through literature survey we can come to know that still there is lot of work to reduce false positive rates and evaluate results with a large database. The methods we are going to implement are given below.

METHODOLOGY

Figure 1 and Figure 2 shows the block diagram of the overall system design for the cancer detection system. The total system is divided into two parts i.e. Training phase & testing Phase respectively as show in above figure. From the block diagram it is evident that SVM is the core for this system. The mammogram image is read which then undergoes image segmentation and enhancement. From the resulting image the required features are collected. These features are then fed to the SVM. This procedure is called as Training. In testing same image processing techniques are used to extract the features. These collected features along with examination result are compared with the available trained features by the SVM classifier. Depending on the comparison result, the classifier gives the result. The proposed method divided into 3 main stages. The first step involves pre-processing, segmentation and filtering procedure. The second step involves feature extraction, and then next and final stage involves classification using SVM classifier.

Image Pre-Processing

Before feature extraction and classification, the input mammogram image is pre-processed as shown in figure 1 in our method 3 steps are carried out in pre-processing. The first step is to convert the RGB image into grayscale image because RGB image takes more processing time. In second step input image is resized to standard size using resize function. Then input mammogram image filtered to remove unwanted noise. Mammograms are medical images that are difficult to interpret, thus a preprocessing phase is needed in order to improve the image quality and make the segmentation results more accurate.

Image Segmentation

Mammogram image segmentation techniques set the focus detecting abnormalities on the region of the breast excluding its background. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. This part plays an important role in processing steps. And if we can get an accurate result in this part, it can help us more in classification scheme. In analyzing mammogram image, it is important to distinguish the suspicious region from its surroundings. The methods used to separate the Region of Interest from the background are usually referred as the segmentation process.

In our method image segmentation or region of interest selection is done by morphological method. Morphological method is the process of picking up a fixed grayscale value and then to classify each image pixel by checking whether it lies above or below this threshold value. Here normalized input image is processed in two levels, first input image is dilated by creating structural element then dilated image is subtracted from normalized input image, maximum value R1 is calculated from resulting image, similarly same procedure continued for level 2 and maximum value R2 is calculated from resulting image. The R2 subtracted from R1 and high intensity pixel is retained based on local area using threshold value, and then it is considered as Region of interest, threshold value is calculated by doing trial and error to 46 mammogram images, the resulting image is considered as segmented image.

Image Filtering

In our method Segmented Mammogram images are then filtered using three different image filters is as shown in figure 1. These filters are intended to help compensate for both intensity variations within an image domain (such as non uniform illumination changes), as well appearance variations between image domains. In our method Difference of Gaussians is utilized to increase the visibility of edges and other detail present in a mammogram image by removing high frequency details including random noise which is present in the image by using the equation (1), which is discussed below. Similarly Gaussian filters also discussed below, here Gaussian filter is used to remove Gaussian noise which is present in the image.

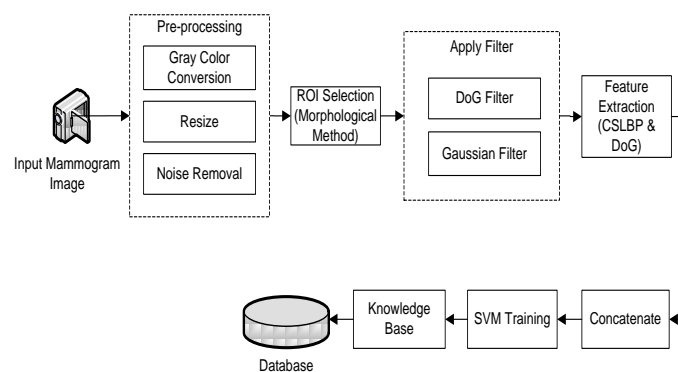


Figure 1: Block Diagram of Proposed Method for Training

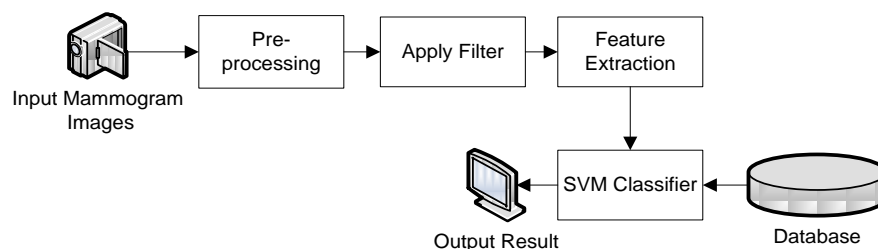


Figure 2: Block Diagram of Proposed Method for Testing

Difference of Gaussian (DoG)

A wide variety of alternative edge sharpening filters operate by enhancing high frequency detail, but because random noise also has a high spatial frequency, many of these sharpening filters tend to enhance noise, the difference of Gaussians algorithm removes high frequency detail that often includes random noise, rendering this approach one of the

most suitable for processing images with a high degree of noise. A difference of Gaussian image is generated by convolving an image with a filter obtained by subtracting a Gaussian filter of width σ_1 from a Gaussian filter of width σ_2 ($\sigma_2 > \sigma_1$). DoGs are linear filters which have been widely used for several vision tasks, including modeling receptive fields in biological vision. We can write the general form of DoG as

$$\text{DoG}(X) = A_1 \frac{1}{2\pi\sigma_1^2} \exp\left[-\frac{\|X\|^2}{2\sigma_1^2}\right] - A_2 \frac{1}{2\pi\sigma_2^2} \exp\left[-\frac{\|X\|^2}{2\sigma_2^2}\right] \quad (1)$$

In practice, the DoG response is identified once the parameters σ_1, σ_2 and the ratio A_1/A_2 are given. In particular, σ_1 can be selected so as to approximately match the size of microcalcification, while, for a fixed σ_1, σ_2 controls the lateral inhibition of the filter.

Gaussian

The Gaussian blur is a type of image-blurring filters that uses a Gaussian function (which also expresses the normal distribution in statistics) for calculating the transformation to apply to each pixel in the image. The equation of a Gaussian function in one dimension is:

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad (2)$$

in two dimensions, it is the product of two such Gaussians, one in each dimension:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3)$$

Where x is the distance from the origin in the horizontal axis, y is the distance from the origin in the vertical axis, and σ is the standard deviation of the Gaussian distribution. The Gaussian smoothing filter has long been used in image processing applications to remove noise contained in high spatial frequencies while retaining the remainder of the signal.

Feature Extraction

In the proposed system once an image is geometrically normalized and filtered using one of the two filters as shown in figure 1, local feature descriptors are extracted from uniformly distributed patches across the mammograms. In this work, the center symmetric Local Binary Patterns features are used. CSLBP features are discussed in D.1. In this method CS-local binary pattern is used for parameter selection by collecting the pixel-wise information from the image see Figure 4 Transforming the input data into the set of features is called feature extraction. Feature is used to denote a piece of information which is relevant for solving the computational task related to a certain application.

Many features have been extracted for the abnormalities of mammograms. The extraction methods of texture feature play very important role in detecting abnormalities of mammograms because of the nature of mammograms. Texture features have been proven to be useful in differentiating masses and normal breast tissues. Texture features are able to isolate normal and abnormal lesion with masses and micro calcification. Feature extraction block diagram is shown in figure 3

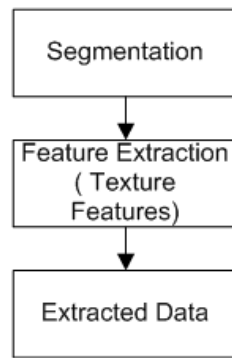


Figure 3: Block Diagram of Feature Extraction

Center Symmetric Local Binary Pattern (CS-LBP)

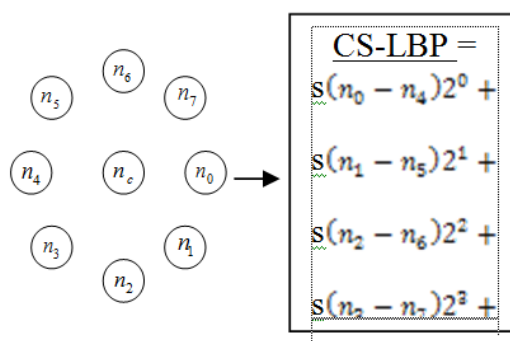


Figure 4: CS-LBP for A Neighborhood of Eight Pixels

CS_LBP is a new texture feature based on the famous LBP operator which has been highlighted successfully for various computer vision problems such as texture classification, face recognition, background subtraction, and recognition of 3D textured surfaces. Instead of describing a center pixel by comparing its neighboring pixels with it in LBP, CS-LBP compares the center-symmetric pairs of pixels, and an example with eight neighbors is shown in Figure 4. The CS-LBP value of a center pixel in pixel (x, y) position is calculated over the neighborhood as follows:

$$CS-LBP_{RNT}(x, y) = \sum_{i=0}^{\left(\frac{N}{2}\right)-1} S(n_i - n_{i + \left(\frac{N}{2}\right)}) 2^i$$

$$S(t) = \begin{cases} 1, & t > T \\ 0 & \text{otherwise} \end{cases}$$
(4)

$$S(t) = \begin{cases} 1, & t > T \\ 0 & \text{otherwise} \end{cases}$$

n_i and $n_{i + (N/2)}$ are the gray values of center-symmetric pairs of pixels of \square equally spaced pixels on a circle with radius \square , and the threshold T is a small value. According to equation (4), the CS- value may be any integer between 0 and $2^{N/2} - 1$. The histogram of the CS-LBP values computed over an image region (the histogram dimension will be $2^{N/2}$) can be used for texture description, and it has been proven to be robust against the changes in illumination. It is also very fast to compute, and do not require many parameters to be set. The value of the threshold T is 1% of the pixel value range in our experiments. Since the region data lies between 0 and 1, T is set to 0.01. The radius is set to 2 and the size of the neighborhood is 8. All the experiments presented in this paper, except the parameter evaluation, are carried out

for these parameters (CS – LBP2, 8, 0.01) which gave the best overall performance for the given test data. It should be noted that the gain of CS-LBP over LBP is not only due to the dimensionality reduction, but also to the fact that the CS-LBP captures better the gradient information than the basic LBP. Experiments with LBP and CS-LBP have shown the benefits of the CS-LBP over the LBP, in particular, significant reduction in dimensionality while preserving distinctiveness.

Different ways of weighting the features are possible. For example, in the case of SIFT, the bins of the gradient orientation histograms are incremented with Gaussian-weighted gradient magnitudes. A comparison of different weighting strategies, including the SIFT-like weighting, showed that simple uniform weighting is the most suitable choice for the CS-LBP features. This is, of course, good news, as it makes our descriptor computationally very simple.

In order to incorporate spatial information into our descriptor, the region is divided into cells with a location grid. Our experiments showed that a Cartesian grid seems to be the most suitable choice. For the experiments presented in this paper, we selected a 4×4 Cartesian grid. For each cell a CS-LBP histogram is built. In order to avoid boundary effects in which the descriptor abruptly changes as a feature shifts from one histogram bin to another, a bilinear interpolation is used to distribute the weight of each feature into adjacent histogram bins. The resulting descriptor is a 3D histogram of CSLBP feature locations and values.

The final descriptor is built by concatenating the feature histograms computed for the cells to form a $(4 \times 4 \times 16)$ 256-dimensional vector. The descriptor is then normalized to unit length. The influence of very large descriptor elements is reduced by thresholding each element to be no larger than 0.2. This means that the distribution of CS-LBP features has greater emphasis than individual large values. Finally, the descriptor is renormalized to unit length.

Difference of Gaussian Features

In this work we proposed new distinctive features called Difference of Gaussian features, this features are extracted from the image by zigzag process. Firstly single level discrete 2D wavelet transform is applied to the mammogram image, it decomposition with respect to either a particular wavelet or particular wavelet decomposition filters that you specify. It computes the approximation coefficients and details coefficients obtained by wavelet decomposition of the input image.

The approximation coefficients are normalized and discrete Fourier transform is computed from the coefficients by using multidimensional fast Fourier transform algorithm, the generated coefficients are then rearranged by moving the zero-frequency component to the center of the array. It is useful for visualizing a Fourier transform with the zero-frequency component in the middle of the spectrum. Then features are extracted in zigzag order to pick the more dominant features.

SVM Classifier

SVM (Support Vector Machine) is a machine learning method that works on the principle of structural risk minimization in order to find the best hyper plane that separates two classes (normal and abnormal). The data used for this SVM is training data and testing data. In this research, testing data are divided into 3 groups. The first group, testing data were taken inside from training data. The second group, testing data were taken outside from training data. And the third group, testing data were taken inside and outside from training data. Grouping is performed to see the accuracy from each group. The process of classification is performed to classify category of normal and abnormal from mammogram image.

The extracted features are finally combined and presented to a Support Vector Machine classifier, Consider the pattern classifier, which uses a hyper plane to separate two classes of patterns based on given examples $\{x(i), y(i)\}$ $i = 1:n$. Where (i) is a vector in the input space $I = R^k$ and $y(i)$ denotes the class index taking value 1 or 0. A support vector machine is a machine learning method that classifies binary classes by finding and using a class boundary the hyper plane maximizing the margin in the given training data. The training data samples along the hyper planes near the class boundary are called support vectors, and the margin is the distance between the support vectors and the class boundary hyper planes. The SVM are based on the concept of decision planes that define decision boundaries. A decision plane is one that separates between assets of objects having different class memberships. SVM is a useful technique for data classification. A classification task usually involves with training and testing data which consists of some data instances. Each instance in the training set contains one “target value” (class labels) and several “attributes”

In the field of medical imaging the relevant application of SVMs is in breast cancer diagnosis. The SVM is the maximum margin hyper plane that lies in some space. The original SVM is a linear classifier. For SVMs [7], using the kernel trick makes the maximum margin hyper plane fit in a feature space. The feature space is a non linear map from the original input space, usually of much higher dimensionality than the original input space. In this way, non linear SVMs can be created. Support vector machines are an innovative approach to constructing learning machines that minimize the generalization error. They are constructed by locating a set of planes that separate two or more classes of data. By construction of these planes, the SVM discovers the boundaries between the input classes; the elements of the input data that define these boundaries are called support vectors.

SVM Training Phase

In data preparation for SVM, various categories for mammogram are prepared. In order to have an accurate classifier we need to collect as much data as possible. This dataset should contain both positive and negative data. In the next step which is scaling data, The SVM algorithm operates on numeric attributes. So we first need to convert the data into numerical format. The original numeric values may be too large or too small in range, thus we have to rescale them to a proper range. To do so, each attribute is scaled linearly to the range of $[-1; +1]$. After scaling the dataset, we have to choose a kernel function for creating the model. For the RBF kernel model, the C and γ parameters have to be set, which are adjusted based on the feature values.

$$d(x) = \sum_{N=1}^{svnum} \alpha y K(X_i, Y_i) + b \quad (5)$$

$$K(X_i, Y_i) = e^{-\frac{\|X_i - Y_i\|^2}{2\sigma^2}} \quad (6)$$

The Linear kernel is the simplest kernel function. It is given by the inner product $\langle x, y \rangle$ plus an optional constant c . Kernel algorithms using a linear kernel are often equivalent to their non-kernel counterparts, i.e. KPCA with linear kernel is the same as standard PCA. The linear kernel model is defined as

$$K(X_i, Y_i) = X_i^T Y_i + C \quad (7)$$

The Rational Quadratic kernel is less computationally intensive than the Gaussian kernel and can be used as an alternative when using the Gaussian becomes too expensive. The quadratic kernel model is given as

$$K(X_i, Y_i) = 1 - \frac{\|X_i - Y_i\|^2}{\|X_i - Y_i\|^2 + C} \quad (8)$$

SVM Evaluation Phase

Using the SVM model, the extracted features will be fed into the SVM system and the normal and abnormal classes in mammograms will be extracted, the SVM classification phase will be executed and texture features will be sent to the SVM model.

The SVM will compare these features with the feature of its entries produced in the training step, and provide the type of class.

RESULTS AND DISCUSSIONS

In this paper, the proposed method includes the input mammogram image pre-proposed as shown in Figure 1 and region of interest calculated in images then filtered with Gaussian filter and difference of Gaussian based on standard deviation and matrix dimensions such as rows and columns. Then the filtered image is used for contrast stretching, and then the features are extracted from the segmented tumor area. The final stage classification is done using SVM classifier.

- SVM has good capacity of generalization.
- SVM is highly robust and work well with images.
- The theory of SVM is well defined and has a very good base of mathematics and statistics.
- Over training problem is less compared to other neural network classifiers.

Thus we have used SVM classifiers to classify the fused feature vector. Implementation is done using MATLAB. For experimentation we have randomly partitioned the dataset training and testing data with the proportion of 70% and 30% respectively. 52 images of two classes are trained in this work, i.e. 26 benign and 26 malignant images respectively. Total 75 images are analyzed as shown in Table 1. The detection accuracy of RBF kernel, non-linear and Linear is calculated using the equation (9), the true positive and true negative values are calculated by confusion matrix as shown in Table 2, Table 3 and Table 4 respectively. Figure 6 shows the performance evaluation plot, it describes the detection accuracy for linear, non linear and RBF kernels which are used in SVM classifier.

$$\text{Accuracy} = \frac{\text{True positive} + \text{True Negative}}{\text{Ground truth Value}} * 100\% \quad (9)$$

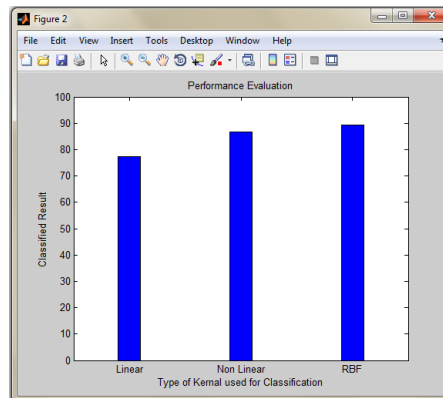


Figure 5: Performance Evaluation

Table 1: Performance Evaluation of the Proposed System

Number of Trained Images	52
Total Number of Images Analyzed	75
Classification Accuracy of RBF Kernel	89.33%
Classification Accuracy of Non-Linear Kernel	86.67%
Classification Accuracy of Linear Kernel	77.33%

Table 2: Confusion Matrix for RBF Kernel

	Malignant	Benign
Malignant	TP (36)	FN (0)
Benign	FP (8)	TN (31)

Table 3: Confusion Matrix for Non-Linear Kernel

	Malignant	Benign
Malignant	TP (30)	FN (6)
Benign	FP (4)	TN (35)

Table 4: Confusion Matrix for Linear Kernel

	Malignant	Benign
Malignant	TP (27)	FN (9)
Benign	FP (8)	TN (31)

CONCLUSIONS AND FUTURE WORK

In this paper we have presented SVM technique for classification of abnormality in digital mammograms and also discussed the CSLBP features as a good tool for features extraction. This research has shown that our method is very effective for the automatic detection and classification of abnormalities in digital mammogram. The evaluation of the system is carried out on standard dataset. The usage of RBF kernel achieved 89.33% accuracy for 75 test images, while linear kernel achieved 77.33% and quadratic kernel achieved 86.67% for same number of images. The proposed method achieves best classification rates SVM from 16 (4×4) sub-images. Also, the SVM classifier gives best classification rate. Our approach divides a ROI image into small regions and computes local texture descriptions using centre symmetric local binary patterns. The combination of these local descriptors in a spatially enhanced histogram provides our final feature descriptor.

The future work focuses on improving the accuracy of early stage cancer detection, this research is still necessary to development and improvement in the system. For the future, is expected to improve segmentation process (find the region of interest) by removing pectoral muscle and removing text noise from digital mammogram. Besides that, system can determine the level of severity (benign and malignant) from classification results. So it can help a doctor to detect and diagnose breast cancer easily.

REFERENCES

1. YM. Abbosh, AF. Yahya, A. Abbosh, (2011). "Neural Networks for the Detection and Localization of Breast Cancer", IEEE International Conference on Communications and Information Technology (ICCIT), Aquaba, pp. 156-159
2. Dave Tahmoush, "Using Image Similarity and Asymmetry to Detect Breast Cancer" Computer Science Department, University of Maryland, College Park.
3. M. A. Duarte, A. V. Alvarenga, C. M. Azevedo, A. F. Infantosi, W. C. A. Pereira, (2011). "Automatic Microcalcifications Segmentation Procedure based on Otsu's Method and Morphological filters" Pan American Health Care Exchanges (PAHCE), ,pp. 102 – 106
4. Y. Ireaneus Anna Rejani, Dr.S.Thamarai Selvi, (2009). "Breast Cancer Detection Using Multilevel Thresholding" Noorul Islam College of Engineering, Tamilnadu, India. g Vol.1(3), 127-130
5. Al. Shamlan, Ali El-Zaart, (2011). "A. Feature extraction values for breast cancer mammography images", IEEE International Conference on Bioinformatics and Biomedical Technology, ,http://en.wikipedia.org/wiki/Image_segmentation pp: 335 – 340
6. Karmilasari, Suryarini Widodo, Lussiana, Matrisyya Hermita, (2011). "Classification of Mammogram Images using Support Vector Machine" Faculty of Computer Science and Information Technology, Gunadarma University, Indonesia.
7. Anupa Maria Sabu, D.Narain Ponraj, Dr.Poongodi, (2012). "Textural Features Based Breast Cancer Detection: A Survey" Dept. of ECE, Karunya University, Coimbatore, India. VOL. 3, NO. 9 Sep
8. Rebecca L. Sawyer, "CS 229 Project Milestone Report: Automated Segmentation of Breast Density".
9. D. S. Y. Ireaneus Anna Rejani, (2009). "Early Detection of Breast Cancer using SVM Classifier Technique," ser. 1, Pages 127-130, Publisher International Journal on Computer Science and Engineering, vol. 1(3).
10. Lowe, D. G. (2004). "Distinctive Image Features from Scale-Invariant Keypoints". Journal of Computer Vision.
11. Yaniv Gal, Andrew Mehnert, Andrew Bradley, Dominic Kennedy, and Stuart Crozier. (2009). "Feature and Classifier Selection for Automatic Classification of Lesions in Dynamic Contrast- Enhanced MRI of the Breast", IEEE Transaction on Digital Image Computing Techniques and Applications, pp 132 – 139.
12. A Lazaro, D. Girbau, and R. Villarino, "Wavelet-Based Breast Tumor Localization Technique Using A Uwb Radar", Department of Electronics, Electrics and Automatics Engineering, Universitat Rovira i Virgili (URV), Av. Paÿsos Catalans 26, Campus Sescelades, 43007 Tarragona, Spain.